



AIAA 2000-2691

**The Modern Design of Experiments:
A Technical and Marketing Framework**

R. DeLoach
NASA Langley Research Center
Hampton, Virginia

**21st AIAA Advanced Measurement Technology
and Ground Testing Conference
19-22 June 2000 / Denver, CO**

For permission to copy or republish, contact the American Institute of Aeronautics and Astronautics
1801 Alexander Bell Drive, Suite 500, Reston, VA 20191

THE MODERN DESIGN OF EXPERIMENTS: A TECHNICAL AND MARKETING FRAMEWORK

R. DeLoach*
NASA Langley Research Center
Hampton, VA 23681-0001

Abstract

A new wind tunnel testing process under development at NASA Langley Research Center, called Modern Design of Experiments (MDOE), differs from conventional wind tunnel testing techniques on a number of levels. Chief among these is that MDOE focuses on the generation of adequate prediction models rather than high-volume data collection. Some cultural issues attached to this and other distinctions between MDOE and conventional wind tunnel testing are addressed in this paper.

Introduction

Wind tunnel research methods that have been entirely adequate throughout the 20th century may now present practical limitations that cannot be surmounted unless these methods are enhanced or extended by the introduction of alternative testing techniques. Increases in measurement precision requirements over time have taken wind tunnel research to a new level, where methods and practices that once served well may now be inadequate for current and future needs.

Measurement precision requirements on the order of one per cent of reading, once regarded as challenging to achieve, are routinely replaced today with requirements an order of magnitude more stringent. To achieve the fractional drag count precision levels demanded in modern aeronautical research requires that all of the variance in a data set acquired by a system as complex and energetic as a modern wind tunnel be explained to within *parts per million*. This is extraordinarily challenging. Regression correlation coefficients as high as 0.9999, once viewed as evidence of an exceptionally

good fit to experimental data, can today be entirely inadequate to meet stated precision goals.

The dramatic increase in precision requirements that has occurred in experimental aeronautics has consequences that are not fully appreciated in all segments of the wind tunnel testing community. For example, subtle interaction effects among independent variables that might once have been ignored without penalty may account for a substantial fraction of today's entire error budget if they are not properly identified and quantified. Interactions exist if a change in some independent variable results in a different response change when a second variable is at one level than when it is at another. For example, if a given change in angle of attack causes a different change in drag when the outboard flap is at one angle of deflection than when it is at another, a two-way interaction exists for drag between angle of attack and flap deflection. If the strength of that interaction is different at low Mach than at high Mach, a three-way interaction exists among Mach, AoA, and flap deflection. Higher-order interactions involving more than three independent variables are also possible. It is not uncommon for subtle high-order interactions to be responsible for unexplained variance in a conventional wind tunnel test that exceeds the entire error budget by over 100%.

Experimental techniques in common use today are described in the literature of formal experiment design as one-factor-at-a-time (OFAT) methods. The name derives from a policy of changing the level of one independent variable (such as angle of attack) sequentially over a specific range, while holding constant all other independent variables (Mach number,

*Senior Research Scientist

Copyright © 2000 by the American Institute of Aeronautics and Astronautics, Inc. No copyright is asserted in the United States under Title 17, U.S. Code. The U.S. Government has a royalty-free license to exercise all rights under the copyright claimed herein for Governmental Purposes. All other rights are reserved by the copyright owner.

control surface configuration, angle of sideslip, etc). Conventional OFAT wind tunnel research methods are essentially blind to interaction effects, which cannot be revealed without systematically considering the effects of multiple independent variable changes.

In addition to interaction effects, assumptions about the stability of sample means have more impact on the quality of research results than ever before. Conventional testing methods require stable sample means as a prerequisite for quality results. That is, while chance variations about a mean value present no serious problem in conventional testing (replication can always drive such errors to acceptably low levels), the mean value itself must be constant over time in order for conventional test methods to generate reliable results. A system that displays this convenient characteristic is said to be in “statistical control”. Systematic variations that ten years ago comprised a negligible fraction of the error budget may with today’s precision requirements consume all of it.

Because of the requirement for stable means in OFAT testing methods, heroic effort and substantial expense are often devoted to characterizing on a continuous basis the state of statistical control in wind tunnels¹. These activities typically involve regularly scheduled tunnel entries with check-standard models, and multiple replicates of a given configuration during production and research operations in order to acquire the data necessary to characterize the long-term and short-term stability of the tunnel.

Unfortunately, statistical control can be an elusive or transient state in practical research applications. A facility pronounced “stable” during last month’s check standard entry may display irksome systematic variations in this month’s research entry. Furthermore, resource limitations render impractical the constant monitoring of system stability throughout a test necessary to guarantee that all research runs are free of systematic variations. At best, only a few such data quality assurance runs can be made. Savvy researchers will concede that uniformly stable sample means are not a realistic expectation for the current state of the art, especially in long tunnel entries. It is necessary to recognize this fact and to alter our research methods to take it into account if we are not to be adversely impacted by instability problems.

A third testing technology issue that has assumed progressively greater importance in recent years is related to the enormous expense of bringing a new commercial jet transport or military aircraft to market. Direct wind tunnel operating expenses, while substantial in absolute terms, are only a small part of

the full cost. Investments in the billions of dollars are required, for which the cost of capital (interest) can range in the millions of dollars per day. It is no wonder that cycle-time reduction is foremost on the minds of aerospace managers today. Unfortunately, management’s concern for cycle time reduction can conflict with the aspirations of wind tunnel research engineers, who tend to regard the relatively time-consuming task of acquiring wind tunnel data in high volume as a necessary prerequisite for reducing inference error probabilities to acceptably low levels. Comfort levels of competent engineers notwithstanding, cycle-time reduction imperatives are likely to be increasingly unforgiving of data acquisition excesses in the future. In such an environment, experimental techniques that include the explicit quantification of data volume requirements for specific technical objectives will be not only desirable, but also essential in a competitive marketplace.

It is the intent of this paper to assert that current wind tunnel testing methods suffer from weaknesses that will cause them to have to be replaced eventually by more robust experimental research methods. Specific weaknesses cited in this introduction include an insensitivity to interaction effects that are key to understanding the underlying physics of flight, an unrealistic requirement for statistical stability that cannot be guaranteed, and a dependence on high volume data collection that will be too expensive to sustain very far into the 21st century.

Fortunately, the importance of interactions among independent variables, the elusiveness of statistical control, and the expense of high-volume data collection are not unique to experimental aeronautics, having been faced squarely and effectively for years by scientific and industrial researchers in numerous other fields. Little else is required for our industry to profit from their experience than to acquaint ourselves with the rich literature that explains their methods, and to consider the practical implementation details associated with bringing them to bear in our specific applications.

If experimental aeronautics is currently capacitated not by measurement technology limitations but simply by our experimental methods, and if solutions to these problems have already been developed in other fields and are readily adaptable to our own, what prevents our industry from embracing these solutions? This paper will make the case that the chief impediments to overcome are cultural, not technical, and will outline certain strategies for doing so. The intent is to cast the cultural issues at least in part in the context of a distinction documented by Clayton Christensen² between what he describes as “sustaining technology”

and “disruptive technology”. Since 1997, Langley Research Center has been developing an alternative wind tunnel testing process to address the limitations of current methods³. Called “Modern Design of Experiments” (MDOE), this method is an example of Christensen’s “disruptive technology”. By recognizing MDOE as a disruptive technology and managing it as such, it may be possible to cope successfully with the cultural issues while continuing to address the specific technical details associated with applications of new experimental methods to the wind tunnel research environment.

A Technical Framework for MDOE

Productivity in late 20th-century experimental aeronautics research is commonly defined in terms of the volume and the quality of *individual data points* acquired in wind tunnel tests. The focus on individual data point quality has resulted in a broad range of process-improvement programs with a common objective of characterizing and ultimately minimizing the error in individual points. Given a source of quality data points, enhancing research productivity is widely believed to be largely a matter of acquiring as many such points as possible in a fixed tunnel entry period. Numerous popular data rate metrics have evolved to quantify trends in productivity defined this way, such as “polars per hour” and “data points per facility per year”. Popular wind tunnel experimental methods are geared toward increasing productivity in this way. For example, the order in which independent variable levels are set is determined largely by data acquisition speed concerns.

The new MDOE process of wind tunnel research introduced at NASA Langley relies heavily on the principles of formal experiment design introduced early in the 20th century by Ronald Fisher⁴, George Box⁵, Cuthbert Daniel⁶, Oscar Kempthorne⁷, and others subsequently extended these methods in a wide range of practical engineering, scientific, and industrial applications. Formal experiment design methods differ in fundamental ways from classical testing procedures in wind tunnel applications. Chief among these is the introduction of an integrated *research* quality and productivity concept that contrasts with the conventional focus on individual *data point* quality, with productivity defined in terms of data acquisition rate.

“Modern” Design of Experiments (MDOE) is a term used to distinguish Langley’s method from conventional OFAT methods, also called “classical” design of experiments in the literature. Within the MDOE framework, quality is defined only in terms of

the precision with which mathematical models can be developed to predict responses of interest (forces, moments, pressures, etc) in terms of independent variables such as angle of attack, Mach number, control surface deflections, and so forth. The wind tunnel researcher’s *raison d’être* is no longer “to get data” (with the implicit focus on high volume and high quality of individual data points.) Rather, the purpose is simply to develop ways and means of predicting response performance.

Parametric models that also yield fundamental insights into the underlying physics of flight may facilitate such predictions, or they may be based on non-parametric models (neural networks, spline fits, wavelets, etc) that rely on numerical methods to provide adequate predictions, especially in highly non-linear regimes. In neither case is the researcher particularly concerned with the volume of wind tunnel data acquired. Once sufficient data are in hand to make response predictions that satisfy documented precision requirements with acceptably low probabilities of inference error (i.e. with acceptably high confidence), there is no particular virtue in acquiring additional data. On the contrary, not only is it not considered highly productive to continue acquiring large volumes of data after sufficient data are in hand to make adequate predictions, such practices are recognized as unnecessary and wasteful (not to say unsophisticated.) Substantial savings in operating costs and in cycle time are often achieved by the simple but effective device of understanding when the experiment is finished, and then stopping.

The response-model focus of formal experiment design also subjugates the traditional goal of localized data point quality to the more global goal of adequate response prediction. With a model capable of adequate response prediction anywhere within a specified range of independent variables, it is no longer necessary to physically set every variable combination of interest. Not only does this often facilitate a substantial reduction in data volume required, it also ensures that the uncertainty of a specific response estimate is not dependent solely on the experimental error of a single data point. Instead, the modeling process facilitates a considerable degree of error cancellation, utilizing the entire ensemble of data to enhance the precision of individual response estimates rather than relying on the quality of individual data points.

Perhaps the greatest strength of formal experiment design in wind tunnel research applications is the effective tactical defense it provides against systematic variations. MDOE practitioners recognize that while statistical control is difficult to sustain in practical research applications, the researcher can himself induce

the attributes of a system in statistical control, by careful attention to the order that data are acquired. Specifically, the randomization of independent variable set-point order insures that response model residuals will be identically and independently distributed about stable means (the defining characteristic of a system in statistical control), even when sample means are unstable. This enables the MDOE practitioner to consistently achieve quality research results, even in the absence of statistical control⁸. The expense and technical difficulty of establishing stable sample means can thus be avoided, by accepting a somewhat reduced data acquisition rate. The sequential ordering of OFAT set points guarantees a higher data rate than can be achieved with MDOE randomization. However, the relatively smaller volume of data typically required in MDOE operations can, and usually does, result in fewer wind-on minutes in spite of the lower acquisition rate.

Figure 1 compares data volume and wind-on minutes for an early MDOE/OFAT comparison conducted during a High Speed Research test in the Langley Unitary Plan Wind Tunnel. Change in wing twist was successfully modeled over a range of angles of attack, Mach, and Reynolds numbers by acquiring only twenty (20) carefully selected data points in the MDOE version of the test, compared to 330 points acquired using conventional OFAT methods to perform the same experiment. (There were actually 660 points acquired in the OFAT version of this test, since each of 330 unique combinations of AoA, Mach, and Reynolds number were set twice to monitor the tunnel's state of statistical control.) The MDOE test matrix called for AoA, Mach, and Reynolds number to all be changed on every data point and set in random order, resulting in a data acquisition rate much less than in the OFAT process. Nonetheless, the MDOE experiment consumed only about 60% of the total wind-on minutes of the OFAT version of the same test, due to the dramatic reduction in data volume.

Figure 2 compares the precision of results obtained via the modern (MDOE) and classical (OFAT) experiment designs for this test. The results are comparable, notwithstanding the fact that the MDOE design required substantially less data.

Figure 3 summarizes the resource savings quantified in a one-year evaluation of MDOE methods. A series of comparison tests were conducted in which MDOE and OFAT versions of the same wind tunnel test were executed back to back. Resource consumption was quantified in both versions of each entry, in terms of wind-on minutes, total data volume, and megawatt hours of electrical power consumed. The comparisons

were conducted in NASA Langley's 14x22 Subsonic Tunnel, 16-Foot Transonic Tunnel, and Unitary Plan supersonic tunnel over a 12-month period beginning January 1997. In all cases the data acquisition rate was much less in the MDOE case due to randomizing the independent variable set point levels to defend against absence of statistical control. However, the reduced data volume requirements of the MDOE method still ensured a net reduction in wind-on minutes that averaged over 60%.

External interest generated during customer observations of the MDOE/OFAT comparison tests lead to a number of additional MDOE tunnel trials, both at Langley and in facilities outside of Langley. As their confidence in the method has grown, the Langley wind tunnel facilities technical staff and management have begun to propose MDOE test segments – and in some cases entire tests – to external customers approaching Langley for access to tunnel time. There have now been thirty-four successful full or partial MDOE tests performed in the last three years. Practical experience in the application of the method is growing, and wind tunnel researchers within Langley and external to Langley have been gaining familiarity with the general concepts. Nonetheless, there exists considerable residual resistance to a wholesale embrace of formal experiment design by the general wind tunnel community.

One of the most frequently identified sources of discomfort among conventional OFAT practitioners on their first exposure to MDOE methods is the MDOE tactic of specifying alterations in the order in which independent variables are set in order to enhance precision. MDOE experiment designs call for independent variables to be set in random order within specific blocks of data. The randomization forces response model residuals to be identically and independently distributed even when the measurement system is not in statistical control, and prevents the confounding that occurs in OFAT testing between independent variable effects and systematic errors when both change systematically with time. (For example, a systematic error that causes early measurements to be biased low and later measurements to be biased high will induce an erroneous counter-clockwise rotation – an increase in slope – in a lift polar if the AoA levels are increased monotonically with time.) The inducing of identically and independently distributed residuals ensures that certain prerequisites for a quality research result are in place, including the fact that standard error computations can be related reliably to confidence interval half-widths. Blocking the data defends against bias shifts that can occur from one block of time to another. For example, scheduling shift changes to

occur on block boundaries prevents the effects of any between-shift differences from contributing to the unexplained variance. If the first shift and the second shift use slightly different procedures, or feature somewhat different skill levels, the effects of these differences on research results can be minimized if the design is blocked properly.

Conventional OFAT practitioners often resist MDOE blocking and randomization tactics for a variety of reasons that include historical factors (“We’ve never done it that way before...”), resource considerations (“This seems to take too much time...”) and simple ignorance (“We always assume no systematic variation...”). Competent OFAT practitioners eventually capitulate as to the *existence* of systematic variations when confronted with evidence of the metronomic regularity with which these effects manifest themselves in wind tunnel data. Hope often persists, however, that the magnitude of these errors will always be small. This justifies an experimental philosophy that couples a belief that a large fraction of bad data points is acceptable as long as they aren’t *too* bad, with a hope that systematic within-test variations will have the good manners never to be excessively large in critical circumstances.

Compelling evidence for these assumptions is rare.

MDOE practitioners generally lean to the conservative view that some tactical defense against possible systematic variations is a good idea, however unlikely such variations may be to occur at any one point in time, and however unlikely it is that they will be serious when they do occur. This relieves the MDOE practitioner from some of the anxieties that may otherwise afflict him. These defense measures have been compared to an insurance policy, which has the value of peace of mind even when no claim is filed.

While OFAT practitioners generally understand and accept the rationale for MDOE experimental tactics intellectually, there remains a visceral resistance the author attributes to complex cultural factors. Those who develop new testing technology must also cope with barriers to the transfer of that technology to potential users. The cultural elements of the MDOE technology development story therefore bear some closer examination.

A Marketing Framework

The model of a marketplace for wind tunnel testing technology, with separate producers and consumers, forms a useful framework for understanding the cultural issues associated with formal experiment design

technology and the changes it represents. In this marketplace as in all others, the consumer dictates how events unfold. In the wind tunnel marketplace the consumers of wind tunnel testing technology are typically the in-house Product Development and Product Support departments of major aerospace firms, while the producers of wind tunnel testing technology are the various R&D departments actually engaged in wind tunnel research. Federal wind tunnel laboratories are also producers of wind tunnel testing technology in this marketplace, whose customers are in-house aeronautical research organizations, and external military and civilian tunnel customers.

The consumer has identified “data” as the principal wind tunnel product of interest. Wind tunnel customers come to wind tunnels to “get data”. Facility operations personnel define their jobs in terms of this task. (Ask anyone associated with wind tunnel operations at NASA Langley what their job is – from the facility managers to the test engineers to the model, data, and operator technicians – and it is almost certain that you will hear something about “getting data”.)

This nearly single-minded focus on *data* has practical consequences. In the competitive marketplace of wind tunnel testing technology, producers recognize the importance of representing data not as a *commodity*, but as a *product*, with certain differentiable qualities. This is an important distinction from a marketing perspective. Commodities are unable to command premium prices in the marketplace because no differentiable features can be assigned to them. The producers of winter wheat in Nebraska cannot command a premium price relative to the producers of winter wheat in Kansas, for example, because winter wheat is simply a commodity that cannot be differentiated according to producer. This forces commodity producers to compete entirely on the basis of price.

To be in a position to stimulate the greatest demand, producers in the wind tunnel marketplace attempt to avoid characterizing wind tunnel data as a simple commodity. They try to differentiate the data available from their facility from the data available elsewhere. This distinction is drawn routinely along two lines – quality and quantity – both dictated by the demands of wind tunnel consumers who drive the process.

The perceived need to “beat the competition” in both quality and quantity drives wind tunnel producers to implement continuous improvement processes in both areas. Continuously improving the product features demanded by your most important customers seems like a certain prerequisite for success. Moreover,

failure to continuously improve along these lines seems like a reliable recipe for eventual disaster. Certainly the wind tunnel producer community at NASA Langley accepts these premises; the quantity and quality improvement processes at Langley are represented schematically by flow charts with entry points but no exits, signifying that data quality and quantity improvement are intended to be ongoing, and to have no end.

Clayton Christensen has highlighted an unfortunate difficulty in the notion of continuous product improvement; namely, that there is an upper limit on the rate that consumers can absorb technology enhancements. In their zeal to ensure that their products meet the most stringent current and future demands of their customers, including especially those in the high-end tiers of their markets, producers run the risk of overshooting the market.

It is the author's belief that producers of wind tunnel testing technology may have in fact overshot the market. Existing technology is capable of far higher precision than is necessary to meet customer requirements when the experiments are properly designed and analyzed. Existing technology is also capable of acquiring data in substantially greater volume than is necessary to support the inferences drawn in a typical wind tunnel test. The MDOE/OFAT head-to-head trials referenced earlier in this paper demonstrated this quantitatively, with MDOE experiment designs achieving equal or better technical results than the corresponding OFAT designs, with substantially fewer data points.

Consumers in the wind tunnel market, unfortunately, have no perception of a current surplus in either the quality or quantity of wind tunnel data. It has not been in the nature of conventional OFAT test planning to quantitatively relate data volume requirements to testing objectives; the methods for doing so are essentially unknown in the OFAT community. Furthermore, tunnel access has been marketed traditionally in units of time, and the conventional OFAT wisdom has been that productivity is measured in terms of total data volume acquired in the fixed tunnel time available. "More" has always been perceived as better, no matter the existing volume of acquired data at any point in time. Indeed, even though MDOE/OFAT comparison studies indicate that OFAT designs specify on average a factor of five times the data volume required to support comparable experimental inferences, consumers in the wind tunnel market continue to demand even higher volume.

"Continuous sweep" operations are an example of demand-driven efforts on the part of wind tunnel test providers to meet the insatiable (if analytically unjustified) demand by their customers for ever higher volumes of data. In continuous sweep operations the sting moves continuously, eliminating the perceived waste of time incurred when the model is paused for a moment to acquire a data point. This also forecloses any option for tactical defenses against time-varying bias errors based on set-point ordering and guarantees an efficient confounding of AoA effects with everything else that changes with time during the test, including all systematic errors. There is a myriad of other technical difficulties also associated with continuous sweep, ranging from the need for precision synchronization of filters in all channels of data to the difficulties in quantifying wall effects with finite-delay pressure tube lengths. It is not the intent of these remarks to argue against "continuous sweep" conceptually; possibly all of the technical implementation problems can be solved satisfactorily, and indeed the fact that this method has been implemented in certain wind tunnels suggests that the problems may have been satisfactorily resolved. The point of these comments is to suggest that the underlying driver for continuous sweep technology development – the perceived need for higher volumes of data that conventional ("pitch-pause") methods provide – is not likely to be scientifically defensible in very many situations. Nonetheless, the demand persists, and tunnel technology producers have little choice but to supply it.

Just as suppliers in the wind tunnel marketplace are compelled by the demands of their customers to overshoot the market in terms of data volume, so too are they driven to overshoot the market in terms of the quality of individual data points. The single-point focus of OFAT testing forecloses certain advantages that accrue when the data set is treated as an integrated ensemble. For example, in MDOE response surface analyses, all data points in excess of the minimum number of degrees of freedom necessary to fit the response model are available to assess the quality of the overall research result. The precision of model predictions increases with the number of residual degrees of freedom, so the MDOE practitioner is not limited entirely by the standard error in individual data points. Furthermore, the MDOE practitioner has at his disposal a number of powerful analytical techniques that can be used to reduce the unexplained variance and thereby increase precision. Blocking, randomization, and replication are all routinely utilized in MDOE testing to improve the quality of research results. Absent these methods, the OFAT practitioner is compelled to devote considerable resources to

perfecting the measurement environment as a prerequisite for achieving quality results. Driven by customer demands for ever-increasing single-point precision, the wind tunnel technology producer is required to commit to complex and expensive “data quality assurance” procedures, including frequent check-standard testing and related activities intended to monitor system stability over various time periods as described earlier.

To summarize the current situation from a wind tunnel market perspective, it is the author’s contention that producers in today’s wind tunnel marketplace supply data with higher precision than is necessary to resolve effects of interest. They also supply data in far greater volume than is necessary to make inferences with satisfactory levels of confidence. It is the essence of what Christensen calls “the innovator’s dilemma” that while these excesses in both single-point precision and data quantity are demand driven and required by customers in this marketplace, those same customers will not continue indefinitely to pay for more than is necessary to meet actual experimental research objectives.

The tradeoffs referenced throughout this paper between quality and quantity in wind tunnel testing may provide an opportunity for MDOE practitioners to achieve a competitive market advantage. Conventional testing technology, with its focus on customer demands for unending increases in data volume and single-point data quality, is an example of what Christensen describes as “sustaining technology”. Sustaining technologies are those that foster improvements in established products, along lines that mainstream customers in major markets have historically valued. There is no question that “data” is perceived as the established product of wind tunnel testing, and that the quality and quantity enhancements with which current wind tunnel technology producers are obsessed is in line with the aspirations of their most important customers. As it has been argued, this is precisely the motivation for these improvements in the first place.

It should be noted in passing that sustaining technology is not limited to “incremental” or “evolutionary” developments. It is possible for a sustaining technology to be radical as well. The “continuous sweep” technology described earlier in this paper is an example of what might be viewed as a radical departure from conventional pitch-pause methods. It is still a “sustaining” technology, however, in that its purpose is to increase data volume, a characteristic of the established wind tunnel product (data) that is already valued and demanded by important customers.

In contrast with sustaining technologies that are focused on data volume and single-point precision, Langley’s MDOE testing process is an example of a *disruptive* technology. Disruptive technologies enjoy little perceived demand early on and in fact they often have *worse* performance than the sustaining technology products with which they compete, at least by conventional metrics. For example, in an environment in which the quality of individual measurements and the total achievable volume of data are prized above all else, MDOE is neutral with respect to single-point data quality, and not only fails to increase data volume, it delivers substantially *less*. On the other hand, disruptive technologies generally appeal early on to a few small segments of the market. In the case of MDOE, the earliest demand was for applications in which either conventional OFAT methods required more data than could be provided, or there were known to be interactions among independent variables that were of interest.

One of the chief barriers to the rapid dissemination of MDOE technology is that it fosters products that are not currently in demand, because they are generally unknown to important customers. For example, MDOE analysis methods are highly objective and quantitative. OFAT analyses are often focussed more on graphical representations of trends, and tend to be more qualitative and subjective. This is not to say that competent OFAT practitioners are unable to make quality inferences using the methods available to them, but the whole MDOE approach to data analysis is much different than what mainstream customers are accustomed to seeing.

MDOE methods are based on the fact that the variance of an ensemble of data points is the key to understanding experimental results. The variance is partitioned into explained and unexplained components. The explained component of the variance is further partitioned in a process that yields insights into the relative importance of main effects and interactions among all the independent variables. The unexplained variance is likewise partitioned, into pure error and lack of fit components. This process permits a very precise evaluation of the quality of the research result, and facilitates scientifically defensible assertions about all inferences drawn from the experiment.

Because mainstream wind tunnel customers have generally never heard of such capabilities, they do not know to ask for them. Likewise, they do not even know how they should respond when someone takes the initiative of offering such products to them, again, simply because of their lack of familiarity. (The typical wind tunnel customer is not likely to know whether to

say “yes” or “no” if someone offers to canonically analyze his ridge phenomena.)

There will be circumstances when MDOE and OFAT methods both satisfy market demands by traditional metrics. For example, high Reynolds number testing in a cryogenic facility such as the National Transonic Facility requires extended periods of time to achieve desired set points because of the time it takes to reduce the temperature as needed to elevate the Reynolds number. A typical configuration study will necessarily feature a small number of configuration changes under such circumstances, and the high-volume aspirations of the customer will not be achievable by any technique. Under such circumstances, decisions about which testing method to use will likely be based on other criteria besides traditional data volume considerations, and the unique features of the MDOE method may assume greater relevance.

Concluding Remarks

The strategy we are following at Langley for introducing MDOE to the general wind tunnel community is a “slow but steady” one, in which all who are interested are offered the opportunity to learn it, and yet none are pressured to accept it. Internally, we have eschewed the choice to “mass-indoctrinate” the entire in-house technical staff in these methods. Instead, opportunities to play substantive early roles in the development of MDOE are reserved for our “best and brightest” young (and young at heart!) engineers. This is partly because of practical resource constraints that make it difficult to involve large numbers of people at once, and also because neither the interest in this new method nor the analytical strengths required to master it are uniformly distributed among all our employees. Nonetheless, all who want to be involved are given the opportunity. The chief prerequisites are enthusiasm, and interest.

MDOE proves the exception to an old saying, in that “familiarity breeds *appreciation*” for this testing method. It has much to offer managers and researchers alike. As time goes by, MDOE will shed its “disruptive” mantle, and the basic tenants of the method will no doubt pass into common knowledge. Formal experiment design offers enough advantages over high-volume data collection as an experimental technique that it is safe to forecast increased use of the method in the years to come. For the time being, though, natural human resistance to change provides an entry barrier for your competitors that you can count on. In a relatively short time, it is likely that formal experiment design methods will be sufficiently common in wind tunnel research that their mastery will offer little

competitive advantage. For the time being, however, there are transient, first-mover advantages available to any who are willing to seize them.

References

- 1) Hemsch, M., et. al. “Langley Wind Tunnel Data Quality Assurance: Check Standard Results”. AIAA 2000-2201. 21st AIAA Advanced Measurement Technology and Ground Testing Conference, Denver, CO. Jun 2000.
- 2) Christensen, Clayton M. (1997). *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Cambridge: Harvard Business School Press.
- 3) DeLoach, R. “Applications of Modern Experiment Design to Wind Tunnel Testing at NASA Langley Research Center”. AIAA 98-0713. 36th Aerospace Sciences Meeting and Exhibit, Reno NV. Jan 1998.
- 4) Fisher, R. A. (1966). *The Design of Experiments*, 8th ed. Edinburgh: Oliver and Boyd.
- 5) Box, G. E. P., and K. B. Wilson (1951). On the experimental attainment of optimum conditions, *J. Roy. Stat. Soc., Ser. B*, **13**, 1.
- 6) Daniel, C. (1976), *Applications of Statistics to Industrial Experimentation*. New York: John Wiley & Sons.
- 7) Kempthorne, O. (1952). *Design and Analysis of Experiments*. New York: Wiley.
- 8) DeLoach, R. "Improved Quality in Aerospace Testing through the Modern Design of Experiments." AIAA 2000-0825. 38th AIAA Aerospace Sciences Meeting and Exhibit. Reno, NV. Jan 2000.

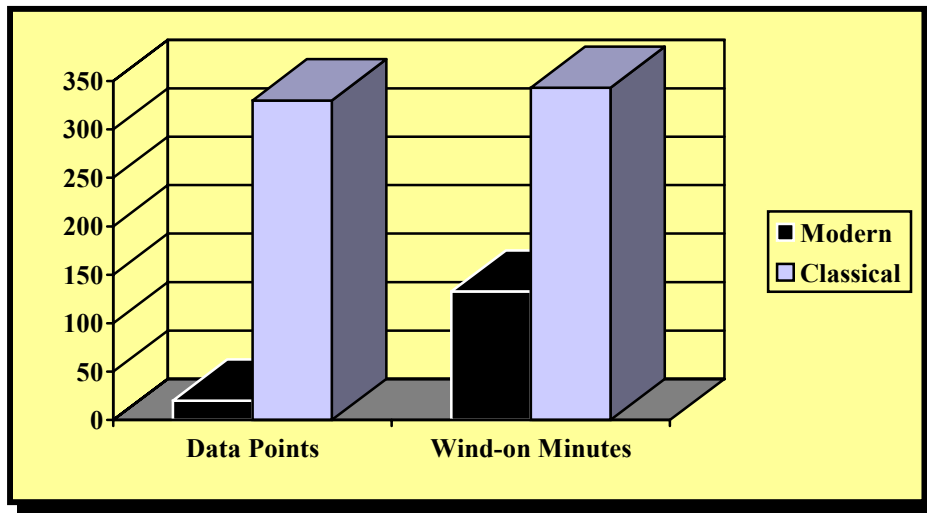


Figure 1: Comparison of Modern and Classical data volume and wind-on minutes

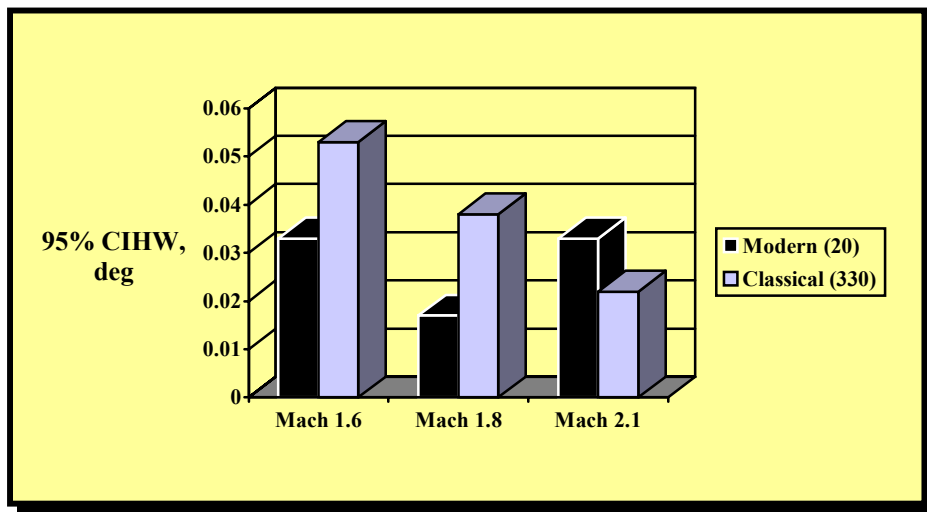


Figure 2: Modern (MDOE) and Classical (OFAT) methods yield comparable precision

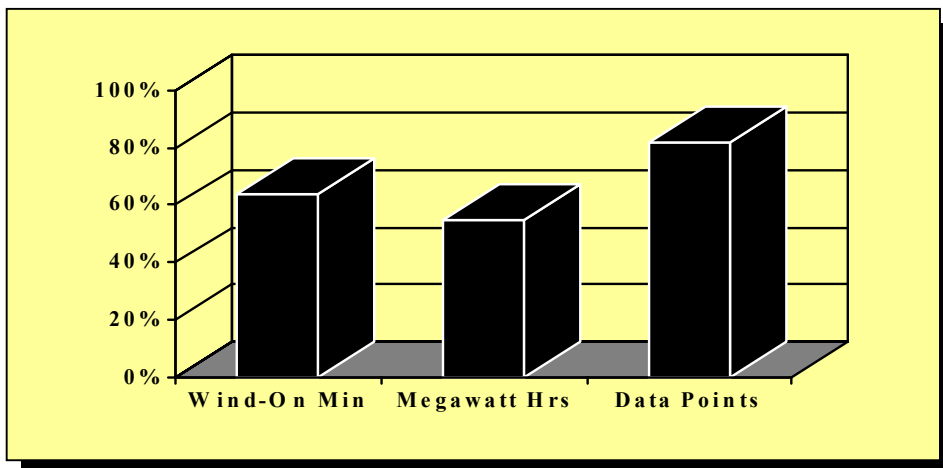


Figure 3: Average MDOE resource savings. Four tests from subsonic to supersonic.